Express Mail Label No.: EL817583554US

METHODS AND SYSTEMS FOR AUCTIONING OF PRE-SELECTED CUSTOMER LISTS

BACKGROUND OF THE INVENTION

[0001] This invention relates generally to financing, servicing loan customers and more specifically to methods and systems for servicing loan customers through dealers of loan products.

[0002] Known customer lead generating systems typically utilize random mailings based on archival data, with little or no attention to trending data or expected incomes. Known modeling solutions are inadequate, because a lifetime probability (a probability for termination in each month in the future of the loan) to prepay loans, described herein as early termination of loans, cannot be accurately computed. Without an accurate lifetime probability, accurate marketing decisions regarding whether the customer is to be offered a promotional or a consolidating offer cannot be made. While customers which would provide high expected incomes for the lender can be sent random mailings with a degree of certainty regarding success of the offer, the calculation of the expected income is only based on a rough approximation. A system without accurate expected income data or lifetime probability for termination does not provide sufficient data for lead development for acquiring new loan business.

BRIEF SUMMARY OF THE INVENTION

[0003] In an exemplary embodiment, a method for operating a system for auctioning customer lists to dealers is provided. The method comprises the steps of prompting a user to select at least one of a plurality of customer profiles, calculating a probability for each customer that the customer will respond to a product offer, grouping customers into distinct lists based upon at least one of the selected profiles and calculated probabilities, and prompting product dealers to bid on the customer lists.

[0004] In a further embodiment, a system configured to generate customer lists for auctioning is provided which comprises a server, a network, and at least one computer connected to the server via the network. The server is configured to prompt a user to select desired customer profiles, calculate a probability for each customer that the customer will respond to a product offer, group customers into distinct lists based upon the selected profiles, and prompt product dealers to bid on the customer lists.

[0005] In another embodiment, a computer is provided which uses at least one of propensity, timing and direct response models to generate customer lists for auctioning. The computer comprises a database of customer data and the computer is programmed to prompt a user to select at least one of a plurality of customer profiles, calculate a probability for each customer that the customer will respond to a product offer, and group customers into distinct lists based upon model output.

[0006] In yet another embodiment, a database is provided which comprises data corresponding to groups of potential customers with distinct behavioral clusters and data corresponding to bids received for each grouping of potential customers.

[0007] In an alternative embodiment, a method for performing an auction of pre-selected customer lists is provided. The method comprises the steps of selecting, from an electronic user interface, at least one grouping of customers, based upon selected profiles, requesting, from the electronic interface, that a list of customers within the groups meeting selected profiles be generated, including a probability of response to an offer, and requesting, from the electronic interface, that registered dealers upload bids for dealer selected groups of potential customers.

[0008] Further, a computer-readable medium is provided which comprises at least one record of customer data, a plurality of rules for matching desired customer profiles to the records of customer data, at least one record

corresponding to a list of customers whose customer data matches the desired profiles, and a record of bids received on said record corresponding to a list of customers.

[0009] Finally, an apparatus is provided which comprises means for prompting a user to select at least one of a plurality of customer profiles, means for calculating a probability for each customer that the customer will respond to a product offer, means for grouping customers into distinct lists based upon the selected profiles and probabilities, and means for prompting product dealers to bid on the customer lists.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] Figure 1 is a system diagram;

[0011] Figure 2 is a block diagram of a network based system;

[0012] Figure 3 is a flowchart diagramming a process for identifying customers with a potential to early terminate an existing loan;

[0013] Figure 4 illustrates a process for identifying potential customers with a propensity to accept an offer;

[0014] Figure 5 is a data diagram for a dealer lead system;

[0015] Figure 6 is a data structure indicating sources of data in a database;

[0016] Figure 7 is a data diagram of a prospect pool auction;

[0017] Figure 8 is a multi-dimensional diagram depicting management of a customer relationship;

[0018] Figure 9 is a data diagram depicting data flows in a customer relationship management system;

[0019] Figure 10 is a user screen used for selecting a listing of accounts;

[0020] Figure 11 is a user screen listing customer accounts and account leads for a number of months;

[0021] Figure 12 is a user screen showing account generation by dealer office;

[0022] Figure 13 is a user screen showing a modeling output of customer prospects;

[0023] Figure 14 is a user screen showing specific information for one customer prospect;

[0024] Figure 15 is a user screen showing results of an offering to potential customers;

[0025] Figure 16 is a user screen tracking results of offers presented to potential customers; and

[0026] Figure 17 is a flowchart depicting a model performance measurement process.

DETAILED DESCRIPTION OF THE INVENTION

[0027] Methods and systems for identifying customers likely to prepay, that is, early terminate their loan contracts before the contracts expire are described below. Customers with a propensity to early terminate are identified using an early termination model, which is configured to predict a probability of early termination for customers based on customer characteristics, for example, age, gender, and income. After probabilities are determined, expected incomes for a lender from a customer are predicted, using the probability information, for a loan and a decision is made whether to offer the customer refinancing or an additional loan offer, in order to retain the customer as a loan customer.

[0028] In addition, methods and systems are described which are configured to use models to determine which of a database of inactive customers have a propensity to make a purchase and the timing of such a purchase. By having an accurate prediction, a dealer of products is able to target the right customer, at the right time, with the right product, before the customer makes their intentions known. The models use as inputs, customer demographic data and other inputs to determine a probability that a future purchase will be made. By offering lists of high probability customers to a dealer of products, a lender is in a position to secure the financing of the products. While stated as an offer to inactive customers, the methods and system described are applicable to those active customers, ready to make additional purchases, for example, the auto loan customers, with a propensity to make a second auto purchase.

[0029] Dealers will pay lenders for generating customer lists which have high probabilities of success in attracting or retain business. A lender generating such a prospect pool, has an opportunity to maximize profit on the customer list by auctioning the list to the dealers of products. Methods and systems for such an auctioning are described.

[0030] The identification and retention of customers is not limited to the customer with a propensity to make purchases which require the taking of a loan. Methods and systems are described herein, where in a retail environment, holders of credit accounts are targeted, in the hope of generating additional purchases, which profit both the retail store and the financial institution through which the store accounts are serviced. By managing the customer relationship through models, again based mostly upon customer demographics, a retailer is able to determine which customers are likely to make only an initial purchase using the store account and those customers who will allow the account to go dormant.

[0031] More specifically, Figure 1 is a block diagram of a system 10 that includes a server sub-system 12, sometimes referred to herein as server 12, and a plurality of customer devices 14 connected to server 12. Computerized modeling and grouping tools, as described below in more detail, are stored in server 12 and can be

accessed by a requester at any one of computers 14. In one embodiment, devices 14 are computers including a web browser, and server 12 is accessible to devices 14 via a network such as an intranet or a wide area network such as the Internet. In an alternative embodiment, devices 14 are servers for a network of customer devices. Computer 14 could be any client system capable of interconnecting to the Internet including a web based digital assistant, a web-based phone or other web-based connectable equipment. In another embodiment, server 12 is configured to accept information over a telephone, for example, at least one of a voice responsive system where a user enters spoken data, or by a menu system where a user enters a data request using the touch keys of a telephone as prompted by server 12.

[0032] Devices 14 are interconnected to the network, such as a local area network (LAN) or a wide area network (WAN), through many interfaces including dial-in-connections, cable modems and high-speed lines. Alternatively, devices 14 are any device capable of interconnecting to a network including a web-based phone or other web-based connectable equipment. Server 12 includes a database server 16 connected to a centralized database 18. In one embodiment, centralized database 18 is stored on database server 16 and is accessed by potential customers at one of customer devices 14 by logging onto server sub-system 12 through one of customer devices 14. In an alternative embodiment centralized database 18 is stored remotely from server 12.

[0033] Figure 2 is a block diagram of a network based system 22. System 22 includes server sub-system 12 and customer devices 14. Server sub-system 12 includes database server 16, an application server 24, a web server 26, a fax server 28, a directory server 30, and a mail server 32. A disk storage unit 34 incorporating a computer-readable medium is coupled to database server 16 and directory server 30. Servers 16, 24, 26, 28, 30, and 32 are coupled in a local area network (LAN) 36. In addition, a system administrator work station 38, a work station 40, and a supervisor work station 42 are coupled to LAN 36. Alternatively, work stations 38, 40, and 42 are coupled to LAN 36 via an Internet link or are connected through an intranet.

[0034] Each work station 38, 40, and 42 is a personal computer including a web browser. Although the functions performed at the work stations typically are illustrated as being performed at respective work stations 38, 40, and 42, such functions can be performed at one of many personal computers coupled to LAN 36. Work stations 38, 40, and 42 are illustrated as being associated with separate functions only to facilitate an understanding of the different types of functions that can be performed by individuals having access to LAN 36.

[0035] Server sub-system 12 is configured to be communicatively coupled to various individuals or employees 44 and to third parties, e.g., customer, 46 via an ISP Internet connection 48. The communication in the exemplary embodiment is illustrated as being performed via the Internet, however, any other wide area network (WAN) type communication can be utilized in other embodiments, i.e., the systems and processes are not limited to being practiced via the Internet. In addition, and rather than a WAN 50, local area network 36 could be used in place of WAN 50.

[0036] In the exemplary embodiment, any employee 44 or customer 46 having a work station 52 can access server sub-system 12. One of customer devices 14 includes a work station 54 located at a remote location. Work stations 52 and 54 are personal computers including a web browser. Also, work stations 52 and 54 are configured to communicate with server sub-system 12. Furthermore, fax server 28 communicates with employees 44 and customers 46 located outside the business entity and any of the remotely located customer systems, including a customer system 56 via a telephone link. Fax server 28 is configured to communicate with other work stations 38, 40, and 42 as well.

[0037] The systems described in Figures 1 and 2 are configured with various models, described below, which are utilized to identify customers likely to early terminate their loan contracts, to attempt to retain the customer with other products. In addition, the models are used to identify high probability customers to which other financial products may be sold, and further, to cluster identified high probability customers into groups. In one embodiment, after clustering, the groups of customers and individual customers are auctioned, for example, based on a probability

determined by the models to dealers of products. Other models have outputs which are utilized to analyze historical retail customer data, and to prepare campaigns to get the retail customer to spend in other areas and through alternative purchasing mediums, for example via the internet. Still other models are configured to determine which inactive customers might be interested in new purchases, for example, auto or retail. Models are implemented as rules within the computer-readable medium of disk storage unit 34.

[0038] In another method utilized to retain customers or attract new customers, a lender offers other loan products, for example, a loan for a new vehicle or a loan for a second vehicle. By cross-selling other loan and loan related products, for example, insurance, auto service contracts, mortgages and any purpose loans, either directly or through a dealer, the lender reduces the chances that the borrower will terminate the existing loan early.

[0039] As another example, at any stage of an existing loan, a customer may be interested in buying an additional vehicle. In one embodiment, using historical data, modeling is used to predict which customers are most likely to respond to an additional vehicle cross-sell, and therefore dealers can target a mailing/telesales cross-sell campaign accordingly. Additionally, some customers will take out a loan for a vehicle and then wish to switch to a leasing arrangement. Analyses performed on the historical data determines which customers are likely to respond to a "loan-to-lease" offer and they are targeted accordingly. An additional way of increasing profitability, is to offer the customer an opportunity to take out a further loan, for example to buy accessories or pay for repairs. Historical data is used to predict which customers are most likely to respond to such an offer, and they are targeted for a pro-active selling campaign.

[0040] Figure 3 is a flowchart 100 depicting a process for identifying potential customers that may be considering an early termination to their existing loan using an early termination model. In addition, the process is utilized in a customer retention program, targeting existing loan customers for a loan cross sell, a refinance of an existing loan or other services as described above in order to retain the customer.

For simplicity, flowchart 100 assumes the existing loans are auto loans, although flowchart 100 should not be considered as being so limited. First, customer data and loan data, for example, income, payment histories and any overpayments, are extracted 102 concerning existing loans and data from earlier loan campaigns 104, for example, customer acceptance or non-acceptance of previous offers, is entered into a loan database 106. Data within database 106 is compared to business exclusion criteria 108, for example, where customer data indicates that a vehicle purchase has taken place within the last six months or is determined to be a cash buyer, to provide a preliminary list of the most promising potential customers. Preliminary potential customers targeted for retention have their credit checked 110 and an expected income from a customer is calculated. Further analysis 112 performed, using customer data, of the customers on the preliminary customer list determines which customers have a high probability to early terminate and which customers have paid off most of their interest payments. After analysis 112, the customer data associated with the preliminary customer list is analyzed using a distribution and return on investment model 114 to ensure that the lender is justified in offering one of a refinance 116 or other loan products 118 and that the customers are the customers likely to respond to a refinancing or alternative product offer. Other loan products 118 typically include a product called an any purpose loan (APL) which is an unsecured cash loan. For the most part, unsecured cash loans do not allow the customer to consolidate their existing loan.

[0041] Early Termination Model

[0042] By borrowing money from other financial sources and prepaying and terminating their contracts, such customers create a serious retention challenge for a lender. In one embodiment, an early termination model is utilized to identify prepaying customers (e.g. early terminators) at least three months before they prepay a loan. The model utilizes six months of loan performance, demographic data (customer information) and rules to predict likely loan terminations. In order to retain these customers, a lender attempts to cross sell the borrower new loans at competitive rates. The annual percentage rate (APR) of the new loan may be lower than the APR

of their existing loan. Although the amount of profit of a refinanced loan, at a lower APR, may be lower than that of the original loan, the sale of the new loan, more often than not, is preferable to losing the customer to a competing lender.

[0043] In the early termination model, calculations of expected income and probability of early termination for the loan are made, using rules which are described below, for each customer, at each stage of the loan. The rules are stored in database 18. Marketing decisions for each individual customer are based on the expected loan income. For example, for customers with either high propensity to early terminate or for customers who have repaid most their interest payments, a new offer is sent out.

[0044] In this approach, the remaining income for each customer at each time (e.g. month) is calculated according to the following rule. If the loan period is T months and the customer is at time t, then the interest payments for all the remaining months, t+1, t+2, …. , T are calculated, and stored within database 18.

[0045] Probabilities to early terminate for t+1, t+2, …. , T are also computed. There are various ways of computing the probabilities of early loan termination for each customer. In one embodiment, a series of logistic regression models are used to predict the probabilities. A_i denotes an event, for example, the customer has early terminated at month i, and A_i^C denotes the complement of event A_i.

[0046] To receive a payment in a given month, the customer has to fail to early terminate in that given month and all previous months. An event that is a customer making a payment in a given month means that the customer has not early terminated for all months up to and including that given month. That is,

$$A_1^C \cap A_2^C \cap ... \cap A_k^C .$$

[0047] By the chain rule formula, $P(A_1^C \cap A_2^C \cap ... \cap A_k^C) =$

$P(A_1^C)P(A_2^C \mid A_1^C)...P(A_k^C \mid A_1^C \cap A_2^C \cap ... A_{k-1}^C)$ and, $P(A_1^C) = 1 - P(A_1)$.

[0048] A_1 , A_2 , A_3 , denotes events where the customer has early terminated in month 1, 2 and 3 respectively. Note that months 1, 2 and 3 represent the months at each 3-month time interval whose probability, $P^{(3)}$, to early terminate is calculated. Then $P^{(3)}$ simply represents the probability of the event by $P^{(3)} = P(A_1 \cup A_2 \cup A_3)$.

[0049] Using probability properties $P(A_1 \cup A_2 \cup A_3) = 1 - P(\ (A_1 \cup A_2 \cup A_3)^C\ ) = 1 - P(A_1^C \cap A_2^C \cap A_3^C) =$ $1 - P(A_1^C)P(A_2^C \mid A_1^C)P(A_3^C \mid A_1^C \cap A_2^C)$ where the events $A_1^C, A_2^C \mid A_1^C, A_3^C \mid A_1^C \cap A_2^C$ represent the probabilities that the customer has not early terminated at month 1, has not early terminated at month 2 given that the customer hasn't early terminated at month 1, etc. The probability calculations are made based upon stored customer information, for example, payment histories, within database 18. Using an assumption that monthly attrition probabilities are equal:

$$P^{(3)} = P(A_1 \cup A_2 \cup A_3) = 1 - (1-p)^3 = p^3 - 3p^2 + 3p .$$

[0050] Having computed monthly interest payments and probabilities for early termination, an expected income from the customer's existing loan is calculated. Rules for calculating an expected income are stored in database 18. For example, if a loan period is T months and the customer is repaying the loan in monthly installments, $I_1, I_2, I_3, \ldots, I_T$ is the interest income, excluding capital repayment, from the customer per month, discounted to the origin of the loan.

[0051] If the customer would stay on book until the end of the loan, (e.g. continue paying interest installments), then the total income would be

$$Total\ Income = \sum_{i=1}^{T} I_i . \qquad \text{(-1-)}$$

[0052] However, if there is a probability that the customer will early terminate at some point in time the expected income is calculated. If P denotes a probability to early terminate over the lifetime of the loan, then the expected total income is $E(Income) = (1-P) * Total\ Income$, where *Total Income* is calculated as above.

[0053] At each point in time (i.e. at every month) using the probability of early termination, $P_i, i = 1,2,...,T$, where T is the loan period, and defined above, the early termination model provides an expected income, given the set of probabilities, as

$$E(Income) = (1-P_1)I_1 + (1-P_1)(1-P_2)I_2 + ... + (1-P_1)...(1-P_T)I_T = \\ = \sum_{i=1}^{T} \left[\prod_{k=1}^{i} (1-P_k) \right] * I_i \qquad (-3-)$$

which is incorporated as a rule within database 18.

[0054] At the end of month m, the expected income for the remaining loan period will be

$$E(Income^{(m)}) = \sum_{i=m}^{T} \left[\prod_{k=1}^{i} (1-P_k) \right] * I_i \,. \qquad (-4-)$$

[0055] Example Calculations

[0056] To clarify the above, below are two example calculations of expected income. The first calculation is at the beginning of the loan period and the second is at an arbitrarily chosen 20th month.

[0057] At the beginning of the loan, (e.g. time = 0 and i=1) expected total income becomes $E(Income) = (1-P_1)I_1 + (1-P_1)(1-P_2)I_2 + ... + \prod_{i=1}^{T} (1-P_i)I_T$.

[0058] At month 20, m=21 and expected income is calculated as $E(Income) = (1-P_{21})I_{21} + (1-P_{22})(1-P_{21})I_{22} + \ldots + \prod_{i=21}^{T}(1-P_i)I_T$.

[0059] At each particular month, m, the set of probabilities is updated by the modeling algorithm where P_i^m is a probability to early terminate at month i, given update information up month m.

[0060] Specifically, $E(Income^{(m)}) = \sum_{i=m}^{T} \left[\prod_{k=1}^{i}(1-P_k^m) \right] * I_i$.

[0061] As described above, monthly interest income calculations $I_i, i=1,2,\ldots,T$, where T is the loan period, are used to compute expected income calculations. Cash flows for the business showing a monthly internal rate of return are calculated for each customer using the following rule, which is stored within database 18: $\sum_{t=1}^{T} \frac{C_t}{(1 + I_{RR})^t} = 0$.

[0062] Knowing the monthly rate of return the monthly installments, M, are decomposed into capital payments, B_t, and interest payments, I_t, i.e. $M = B_t + I_t$. In one embodiment, M the monthly installments are constant during the lifetime of the agreement. The model is configurable for variable monthly installments as well.

[0063] As an illustration, a decomposition for the first two months is given by $I_1 = I_{RR} * L \Rightarrow B_1 = M - I_1 \Rightarrow L_1 = L - B_1$, for month one and $I_2 = I_{RR} * L_1 \Rightarrow B_2 = M - I_2 \Rightarrow L_2 = L - B_2$ for month two, where L_1, denotes the remaining capital amount, and L denotes the original loan amount.

[0064] Losses are avoided by not targeting all consumers. To facilitate loss avoidance, again using the auto loan as an example, an auto finance income curve, also described as a yield curve, is used to calculate an income a lender

can expect to receive from the customers for the remaining months of the auto loan. Yield curves are well known in the art and are extensively utilized in financial and government sector in computing yields for various financial instruments. For example, yield curves on bonds are used to show the yields for different bond maturities and also to show the best values in terms of financial growth. In one embodiment, the early termination model is linked with the yield curve within system 22 to identify which consumers should be solicited for loan cross-sell before they terminate their current loan insuring that the lender does not leave any funds not working for them, while also addressing customer retention.

[0065] Generation of Customer Leads for Dealers

[0066] In addition to the possibility of refinancing and the offering of other loans to try to prevent early loan terminations, other models for customer life cycle management are embodied as rules to identify appropriate communications and actions to present to a finance customer throughout his/her finance lifecycle. Still other models are used to determine which inactive customers, those without current loans, to target for various products, for example, direct auto loans to consumers to buy a new car, buy insurance (including, but not limited to, payment protection insurance (insurance against becoming unable to make payments because of illness), personal accident insurance, and auto insurance), personal loans, and debt consolidation loans to repay auto loans, credit card debt and other non-lender financed loans.

[0067] Figure 4 illustrates a process for identifying inactive customers who would likely be interested in a new automobile purchase. Such a process is capable of being hosted on a system such as system 22 (shown in Figure 2). Models are combined and utilized to predict when and which customers will purchase direct loans. An inactive customer database 120 contains records of inactive former customers. It should be noted that database 120 is also described as database 18 (shown in Figure 1) or database 34 (shown in Figure 2). The customer data is modeled using propensity and timing models 122 embodied as rules within database 120, to determine a customer group 124 who have a propensity to purchase and when

a probable time of purchase for those customers will occur. A subset of group 124, is modeled using a direct response model 126 to increase response and conversion rates and outputs a group 128 likely to respond to a campaign. A campaign 130 targeted to group 128 is initiated and feedback 132 regarding purchases and non-purchases are input into database 120. Such modeling reduces mailing costs by implementing period-specific customer targeting.

[0068] Propensity and timing models 122 use customer characteristics, including, but not limited to, age of current auto, earnings, probable loan period, payment history, percent of current auto that was financed, and current installment payments, to predict a buying period and propensity to purchase for each customer. Predictions of timing and propensity are determined within the models based on computed probabilities, as described above in the early termination model. Direct response model 126 supplements propensity and timing model to increase response and conversion rates using characteristics including, but not limited to, earnings, a time amount the customers has had loans, a time amount the customer has been a customer, marital status, and percentage of current auto that was financed, again based upon computed probabilities. The process described in Figure 5 is further configurable to predict when a customer might purchase a second auto. Other customer characteristics that may be utilized in the model include price of auto, genders, arrears, interest charged, former purchase was a new vehicle, customer age, value of trade in auto, time in employment, credit rating, customer home owner, how long since customer, when last proposal was made and total amount repaid from first loan. Although campaign 130 is described as a mailing campaign, targeted customers can be contacted by one or more of e-mail, fax, cell phone, and telephone. Data gathered in database 120 comes from multiple sources, for example, Internet, legacy data, fax, phone and cell phone.

[0069] In another embodiment, customer characteristics are clustered, using rules stored within database 120, to segment consumers and identify sales opportunities within active and inactive customers. Examples of clustered customer groups, for example, would include young used-car buyers, bargain hunters,

families, high spenders and reluctant borrowers. Clustering is based upon groupings of customer characteristics, the characteristics being the customer demographics utilized as model inputs, listed above.

[0070] As described above, customer information and model predictions regarding customer purchases are stored in database 120 and are accessible interactively at any time via the internet. The above system and modeling descriptions provide suggested triggers which activate a cross-sell process for an account, generate leads for increased sales, and predict which customers are likely to terminate early or respond to a further marketing campaign.

[0071] In one embodiment, the process shown in Figure 5 is utilized to generate leads to supply to dealers for the sale of new products. As described above, system 10 uses modeling as described herein to identify customers from both active as well as inactive files who are likely to respond to an offer and make purchase a product. Since, in one embodiment, system 10 is web-based, dealers or other industries are given access to system, for example, via the internet, to access customer leads. Customers that have been contacted may also access the Internet to view offers instead of visiting the dealership.

[0072] Figure 5 is a data diagram 150 for such a dealer lead system. In one embodiment, dealer owned databases 152 of customer data are combined with the customer data in database 120, which is lender owned (shown in Figure 4) and modeled using system 22 according to the process described in Figure 4. By combining data within dealer owned databases 152 and data within database 120, dealer specific customer contact lists are generated, using models such as those described above, which are stored in a dealer customer database 158. System 22 (shown in Figure 2) is configured to prevent databases 152 owned by one dealer from being co-mingled with another dealer's database 152.

[0073] Figure 6 is a data structure 170 indicating sources of data to be stored within a dealer database 152 (shown in Figure 5) including, but not limited to, web 172 (or Internet), legacy data 174 from previous sales, call center 176 where

prospective customers call a dealer and in turn, the dealer receives customer data, fax 178 and phone 180 are other sources where a dealer can store data within database 152. By modeling with detailed customer data, as described above, including clustering of customer groups as described above, dealers provide better service to potential customers by identifying specific needs for each potential customer. Customer data updates uploaded to database 120 (shown in Figure 4) can cause an automatic update to modeling outputs, since system 22 (shown in Figure 2) is web based and can be configured to provide such automatic updates.

[0074] Auctioning of Customer Leads to Dealers

[0075] As described above, the modeling methods described herein pro-actively seek prospective borrowers from the wider population and can therefore identify prospective customers with a propensity to buy, and further identify the timing of such customer purchases. As also described above, dealers with such customer information would typically market to these prospective customers to invite them to buy their products and take a loan through the dealer from the lender that has provided the modeled prospective customer list. Since the modeled prospective customer list is valuable to dealers wishing to make sales, dealers will pay a lender who provides a list of high probability prospective customers.

[0076] Figure 7 is a data diagram 200 diagramming a prospect pool auction being conducted by a lender. Lender owned databases 202 with data sources as described above and others are entered into system 22 which incorporates the modeling processes described above. Dealers 206 bid for the pools of prospects for purchase of their products, via an auctioning 208 process. System 22 (shown in detail in Figure 2) is configured to conduct the auction process and, as described above, generate customer lists for auctioning via modeling and clustering processes. System 22 includes rules to calculate for each customer a probability that a customer will respond and make a purchase decision once a product offer is made to that customer based on the modeling process. The auction is a probabilistic auctioning. Participants in the auction have access to depersonalized information (no names and addresses) available on the customers as well as the probability of response to an offer (i.e., car

loan, personal loan). Dealer 206 participants can bid on each customer one by one or they can bid on a group of customers who belong to a particular cluster. Dealer participants have capability to set a range of probability of response, for example, 90 percent and higher, they are interested in and bid on prospective customers who fall within that range. Bids, customer probabilities, and clustered customer groups are stored in database 120, as are rules for matching the bids with the probabilities and the customer groupings.

[0077] The auction described herein is different from a typical auction where the owner of an asset offers the asset for auctioning. The auction described is also different from the reverse auction, which is sometimes called a demand driven auction, where bidders who express their willingness to pay a set amount of dollars for a given service and then offers that amount to suppliers of those services who are willing to meet the price offered by consumers. In essence, demand is driving the auction. The auction described herein is based on an early identification of the needs of customers for a particular product before customers express those needs or demands. Customers' needs are identified by modeling customer data within database 120 (shown in Figure 4), and the dealer and lender do not have to wait for potential customers to express their demand for services. The lender auctions pre-selected customer lists to businesses (i.e., dealers) who in turn lets the lender finance the purchase of the products for the customers.

[0078] Although applicable to any type of loan sale, the examples and models described herein use auto loans only as an example, the scope of modeling applications herein described should not be construed to be so limited.

[0079] Customer Relationship Management

[0080] The modeling and identification of customers who have a probability of early loan termination or a propensity to accept a financing offer have thus far been described, including identification of high probability new customers in the form of lists for access and auction. However, all of the customer list generation has been described in terms of the sale of loan products to these customers. Such

modeling and identification techniques are applicable to retailers of products, for example, department stores, where capturing store account financing is one goal of a lender. Retailers are engaged in dynamically competitive businesses, for the most part, and are faced with reducing margins due to competition. Such retailers also have a need for proactive, agile marketing.

[0081] Therefore modeling solutions are applicable for targeting of retail customers, providing such modeling solutions are able to deliver specific, targeted customer relationship management solutions. Such solutions enable retailers to increase customer expenditures and enhance loyalty using data driven decisions from the models to target the right communications at the right times to the right customers.

[0082] In addition to timing, propensity and direct response modeling as described above, a retail customer relationship management solution includes a hit and run model and a dormancy model. Hit and run customers do not use their store accounts after an introductory purchase or period. One example is where a discount is offered for a first purchase, provided that purchase is done on a store account. Persuading a number of such customers to make another purchase on the store account is, of course, profitable for the store and for the lender who services the store accounts.

[0083] A hit and run model includes application variables, transaction variables and geo-demographic variables. Application variables include, but are not limited to, gender, zip code, date of birth, date store account opened, payment methods, card protection information, availability of telephone, and holder of additional store account. Transaction variables include merchandise groups of first and last purchases, division of first and last purchases, total value of all transactions, value spent on first and last days, total number of transactions average transaction value, average daily transaction value number of different divisions, number of different transaction days, largest and smallest daily transaction values. Geo-demographic variables include zip code data, branch data (which stores purchases were made), and distance data from the customer's home to their preferred branch.

[0084] The hit and run model is configured to predict a propensity to hit and run using account data from a first week of an accounts existence. Propensity to hit and run may also be determined based on particular transaction variables. The hit and run model and the dormancy model, described below, use customer demographics and account and spending data for determining a probability of hitting and running or dormancy, respectively, similar to the process described above for the early termination model.

[0085] A dormancy model is similar to the hit and run model, the difference in definition being that there were no introductory purchases made. Transaction variables in a dormancy model include, but are not limited to, gender, town, customer age, account age, number of purchases in 12 months, value of purchases over 12 months, number of total purchases, value of total purchases, value of transaction, card style, division, style, number of total purchases/account age, value of total purchases/account age, and value of total purchases/number of total purchases.

[0086] Customer relationship management solutions are based on key performance indicators, such as the variables described above, and an optimization of spending patterns along retail dimensions through actionable analysis. Solutions include access to embedded clustering analytics provided by models, which are embodied as rules in database 18, for example, to further profile cluster groups, for example, young low spenders, shopaholics and high-frequency shoppers, against key performance indicators, for example, amount spent and transaction frequencies. Such an application of modeling rules provides ad-hoc views of data to be reconstructed and allows profiling of cluster groups against a listing of attributes, for example, age, gender and geographic area.

[0087] In one embodiment, the models are embodied as rules within database 18 to predict future spending of each customer in a database in a specified time period. In another embodiment, models are configured to rank order customer accounts that have not had spending activity in a given number of days, to predict a likelihood of spending in the next given number of days. Weeks and months are time frames which are also implemented in the models. In still another embodiment,

models are embodied as rules to rank order customer accounts that have had spending activity only once, based on a probability of there ever being any future activity within the account.

[0088] The models are used with planned promotional events, configured to determine an optimal mailing list and a size of the mailing, based on a likelihood of response, an overall response rate and profitability margin. Using such models, product purchase patterns are identifiable and key variables which will indicate trends are also predictable. Once a retailer or system user has entered available information into databases, a one-click marketing solution to a campaign is available based on model output. Customer relationships to be managed, include, for example, store credit cards, bank cards, financing of sales, and customer loyalty. Those customer relationships are channeled through multiple retail mediums including point of sale, home shopping, E-commerce and Digital television.

[0089] Figure 8 is a multi-dimensional diagram 220 depicting management of a customer relationship. Although the example in Figure 8 is exemplary, the example is not exhaustive. Diagram 220 is three dimensional, depicting retail relationships with customers based on value 222, frequency 224, and departments previously visited 226. The customer relationship is further managed by identifying departments not typically visited, for example, by segmentation. Such segmentation is used to prepare marketing solutions which expose the customer to those non-visited departments in an attempt to increase sales frequency and visits to those departments.

[0090] Figure 9 is a data diagram 240 depicting data flows in an exemplary customer relationship management system. Data diagram 240 includes a number of databases 242 which are populated with customer retail store card data 244 and the model variables as listed above. Data, stored in databases 242 are subjected to modeling analysis 246, for dormancy, hitting and running, clustering, attrition, spending behavior, response probability, and cross-selling propensity, and further subjected to, in the embodiment shown, a distribution model 248 and return on investment model 250. The customer data is then ranked 252, based on the return on

investment model 250, and made available to retailers as retail store card data 244. In one embodiment, the dormancy and hit and run models are implemented using a system, for example, a system such as system 22 (shown in Figure 2) enabling retailers access websites 254 to view model outputs and therefore construct customer campaigns, for example, websites 256 with personalized offers for targeted customers.

[0091] Application of historical data allows the building of models to attempt to capitalize on the future demand of customers. The hit and run and dormancy models described above are configurable to take into account seasonality, socio-economic trends, information on past levels of demand, and other factors enabling dealers to optimize their capacity by predicting those accounts holders who will go dormant.

[0092] Screen Shots

[0093] Figure 10 is a user screen 260 displayed by system 10 or system 22 which is used for selecting a listing of accounts. Screen 260 is configured for selecting customer groupings for identification of early terminating accounts, identifying customers with a propensity to take out a loan for purchase, or to manage customers with retail accounts. As shown in screen 260, a user can select a listing of total accounts and balances, new accounts and balances and a report showing dealer office performance. It is contemplated that screen 260 also include a selection for inactive accounts.

[0094] Figure 11 is a user screen 262 displayed by system 10 or system 22 listing customer accounts and account leads for a number of months. Through screen 262 a user is able to track a number of accounts that have resulted in leads for customer retention, as described by the early termination model, or that have resulted in a customer contact with a dealer regarding possible purchases, resulting in new or additional loan activity for a lender.

[0095] Figure 12 is a user screen 264 displayed by system 10 or system 22 showing account generation by dealer office. As shown in screen 264,

dealers are ranked according to performance indicators. Such performance indicators include value of units sold or refinanced and number of units sold or refinanced. The reports can be generated for user selectable time periods (e.g. year, month).

[0096] Figure 13 is a user screen 266 displayed by system 10 or system 22 showing a modeling output of customer prospects. Screen 266 is used in both customer retention programs (e.g. identification of early terminating customers) and potential or repeat customer identification. As shown in screen 266, for a number of customers contacted, there is a number of those customers which are interested, moved, responded, accepted or converted to a new finance opportunity. Interested customers are ranked according to probabilities to refinance (or purchase new products). Listed customers are selectable for a display of detailed customer information (shown in Figure 14).

[0097] Figure 14 is a user screen 268 displayed by system 10 or system 22 showing detailed information for one customer prospect, the prospect being selected from screen 266 (shown in Figure 13). In screen 268 customer information includes name, address, age, interest rate of active loans, balance of loans, marital status, age, gender, if homeowner, vehicle age and vehicle price. Further fields indicate what cluster the customer is in, including a description of the cluster, for example, reluctant borrowers. Further included in screen 268 are fields indicating if the customer has been contacted, if the customer is interested, and what the result of the contact is: accepted, responded or converted.

[0098] Figure 15 is a user screen 270 displayed by system 10 or system 22 showing results of offerings to potential customers. Screen 270 indicates, by cluster, a number of accounts, a number responded, a number accepted and a number converted. Screen 270 further includes a graph indicating deal measurements, for example, acceptance, by clustered customer groups.

[0099] Figure 16 is a user screen 272 displayed by system 10 or system 22 tracking results of offers presented to potential customers by volume. As shown on screen 272 volume metrics indicate a number of accounts that have

responded, accepted, and converted by week. Rate metrics indicate a response rate, net conversion rate and an acceptance rate by week.

[00100] Model Accuracy

[00101] Figure 17 is a flowchart depicting a model performance process 300. In a characteristics analysis 302, each predictive characteristic calculated from observation data is reviewed to success of the characteristic in predicting risks. In a optimal banding process 304, an entropy measure is utilized to optimally split continuous variables into different bands. A correlation analysis is used 306 to identify highly correlated variables, such as those used in the models described above. Where a degree of correlation is not acceptable, the variable pairs are considered for exclusion and analysis by testing each of the two variables in turn for correlation with the outcome variable. The explanatory variable with the highest correlation with the outcome variable is retained for model development, and the other explanatory variable is excluded.

[00102] Logistic regression models are built 308 for each time period and the accuracy of the models is then tested with the validation set, for each time period. The models are validated 310 using a set of Lorenz curves, not shown, as a measure of the predictive power of the models. Logistic regression is a form of statistical modeling that is often appropriate for dichotomous outcomes, for example good and bad. Logistic regression describes the relationship between dichotomous variable and a set of explanatory variables. A logistic model for an i^{th} experimental unit is defined as:

$$\Pi_i = F(\beta_0 + \sum_{j=1}^{J} \beta_j z_{ji}) = \frac{\exp\left(\beta_0 + \sum_{j=1}^{J} \beta_j z_{ji}\right)}{1 + \exp\left(\beta_0 + \sum_{j=1}^{J} \beta_j z_{ji}\right)}$$

where $F(\beta_0 + \sum_{j=1}^{J} \beta_j z_{ji})$ is a cumulative density function for the logistic distribution, and can also be expressed as a linear function of parameters:

$$\ln\left(\frac{\Pi_i}{1-\Pi_i}\right) = \beta_0 + \sum_{j=1}^{J} \beta_j z_{ji} .$$

[00103] The expected value and variance of the logistic regression model are 0 and $\frac{\Pi^2}{3}$, respectively. One of the advantages of statistical modeling is that measures of association are functions of parameters. An association is said to exist between two variables if the behavior of one variable is different, depending on the level of the second variable. In one embodiment, associations are evaluated within a matrix using a statistical measure called concordance. Measures of association are based on the classification of all possible pairs of subjects as concordant pairs. If a pair is concordant, then the subject ranking higher on the row variable also ranks higher on the column variable.

$$C = \sum_{i=1}^{I} \sum_{j=1}^{J} X_{ij} \left[\sum_{k>i} \sum_{l>j} X_{kl} + \sum_{k<i} \sum_{l<j} X_{kl} \right]$$

where X*ij* stands for the number of observations for i^{th} row and j^{th} column.

[00104] Similarly, if a pair is disconcordant, then the subject ranking is higher on the row variable and lower on the column variable.

$$D = \sum_{i=1}^{I} \sum_{j=1}^{J} X_{ij} \left[\sum_{k>i} \sum_{l<j} X_{kl} + \sum_{k<i} \sum_{l>j} X_{kl} \right]$$

[00105] Also, the pair can be tied on the row and column variable.

$$T = \sum_{i=1}^{I} X_{ii} \left[\sum_{k=j\neq i} X_{kk} \right]$$

[00106] The higher the concordance, the larger the separation of scores between good and bad accounts. The concordance ratio is a non-negative number, which theoretically may lie between 0 and 1. However, in practice most scores are between 0.6 to 0.95. To ensure that the parameter estimates β_k in the logistic regression model have comparable magnitudes for the different independent variables x_k, which have different units (i.e. age, deposit, value of the car, interest rate, etc) all the continuous variables have been standardized, i.e.

$$\widetilde{X} = \frac{X - E(X)}{\sigma(X)}$$

where $E(X)$ is the mean of X and $\sigma(X)$ is the standard deviation of X. Therefore by looking at parameter estimates of the models, magnitudes of different independent variables are compared.

[00107] Information collected on customers for modeling input is also used to service more than the financial industry. For example, in a travel industry application, information on the use and servicing of the car is used to predict when long trips will be made, presenting opportunities to cross-sell products relating to the travel industry (e.g. travel insurance, foreign currency).

[00108] In addition, some customers will want to take advantage of an offer for a product upgrade or replacement. For example, in the auto industry, many auto dealerships offer schemes whereby they will upgrade a customer's vehicle at some point in the future, taking the original vehicle as a part of the exchange. The models described herein make it possible to estimate the residual value of the original vehicle as a function of time, and therefore predict the residual value of the vehicle at the time that it is accepted as a trade-in, knowledge useful to both customers and dealers.

[00109] Modeling allows dealers to optimize their portfolio by taking into account particular risks that the dealers may wish to minimize (e.g. a change in the delinquency rate). Modeling finds the correct decision-making process to make

with regard to each individual customer, with respect to the type of loan they should be offered, so as to achieve a portfolio which maximizes profit subject to customer-defined risk constraints.

[00110] The models are further used to examine the profitability of an individual client at present and into the future. By taking into account loans in the loan portfolio, expected delinquency of accounts, expected early termination, changes in interest rates, etc, to give a complete picture of the revenues and costs of the client. Such a process allows application of individual processes (e.g. cross-selling a personal loan to a customer. The combination of the processes build a portfolio that a dealer may apply selectively to the customer base over the lifetime of the loan.

[00111] In addition, web-enabling of the processes described herein is contemplated so that dealer are able to use models over the internet to model or analyze the customer base in their own way, and will have an instant, interactive tool, to indicate which processes should be applied to which customers. Further, development of applications via wireless application protocols (WAP) so that dealers are able to access and apply analyses from the models without the requirement of access to a telephone line or other medium is foreseen. Traditional marketing techniques as postal mailings, couriers as well as Electronic mail targeting are enhanced by targeting the right customers using the modeling outputs.

[00112] Modeling processes occur at certain points in the customer lifecycle. Using these processes a mechanism is enabled by which different processes are applied to the customer at each stage of the loan for maximum profitability. While the invention has been described in terms of various specific embodiments, those skilled in the art will recognize that the invention can be practiced with modification within the spirit and scope of the claims.